umoore

1652

#10
worm

PAGE: 1   **RAW SEQUENCE LISTING**   DATE: 02/01/2000
                       PATENT APPLICATION US/09/147,947A   TIME: 14:05:12

Input Set: **I147947A.RAW**

> **This Raw Listing contains the General Information Section and up to first 5 pages.**

ENTERED

```
  1   <110> APPLICANT: TSURUOKA, Nobuo
  2         YAMASHIRO, Kyoko
  3         YAMAGUCHI, Nozomi
  4   <120> TITLE OF INVENTION: Novel Serine Protease
  5   <130> FILE REFERENCE: 001560-349
  6   <140> CURRENT APPLICATION NUMBER: US/09/147,947A
  7   <141> CURRENT FILING DATE: 1997-03-24
  8   <150> EARLIER APPLICATION NUMBER: PCT/JP98/03324
  9   <151> EARLIER FILING DATE: 1998-07-24
 10   <150> EARLIER APPLICATION NUMBER: JP 9/213969
 11   <151> EARLIER FILING DATE: 1997-07-24
 12   <160> NUMBER OF SEQ ID NOS: 6
 13   <170> SOFTWARE: PatentIn Ver. 2.0
 14   <210> SEQ ID NO 1
 15   <211> LENGTH: 20
 16   <212> TYPE: DNA
 17   <213> ORGANISM: Artificial Sequence
 18   <220> FEATURE:
 19   <223> OTHER INFORMATION: Description of Artificial Sequence:Synthetic DNA
 20   <220> FEATURE:
 21   <221> NAME/KEY: misc_feature
 22   <222> LOCATION: (9)..(12)
 23   <223> OTHER INFORMATION: Nucleotides 9 and 12 are n wherein n = any
 24         nucleotide.
 25   <400> SEQUENCE: 1
 26         gtgctcacng cngcbcaytg                                             20
 27   <210> SEQ ID NO 2
 28   <211> LENGTH: 20
 29   <212> TYPE: DNA
 30   <213> ORGANISM: Artificial Sequence
 31   <220> FEATURE:
 32   <223> OTHER INFORMATION: Description of Artificial Sequence:Synthetic DNA
 33   <220> FEATURE:
 34   <221> NAME/KEY: misc_feature
 35   <222> LOCATION: (6)..(9)
 36   <223> OTHER INFORMATION: Nucleotides 6 and 9 are n wherein n = any
 37         nucleotide.
 38   <400> SEQUENCE: 2
 39         agcggnccnc cdgartcvcc                                             20
 40   <210> SEQ ID NO 3
 41   <211> LENGTH: 2614
 42   <212> TYPE: DNA
 43   <213> ORGANISM: Mouse
 44   <220> FEATURE:
```

W--> OK (line 26)

W--> OK (line 39)

45   <223> OTHER INFORMATION:
46   <400> SEQUENCE: 3
47   cgagggtggg gtggaggtcg gactccgggc tacagagctc ctggcgctca tcgcctctgg   60
48   ctccagcctt tgcttcgcgg ggctgaccct ttgggtcccg gtgtgatcct ccagctgccc   120
49   cgggggctgg gacagcaggg cggcggcgcg agcgtgggag ggggctctag gactctgccg   180
50   gccccgcccc gccccctccg cggggacccg gagcccagca tggaccacac tcggcgccgc   240
51   agcc atg gcg ctc gcc cgc tgc gtg ctg gct gtg att tta ggg gca ctg   289
52   Met Ala Leu Ala Arg Cys Val Leu Ala Val Ile Leu Gly Ala Leu
53   1   5   10   15
54   tct gta gtg gcc cgc gct gat ccg gtc tcg cgc tct ccc ctt cac cgc   337
55   Ser Val Val Ala Arg Ala Asp Pro Val Ser Arg Ser Pro Leu His Arg
56   20   25   30
57   ccg cat ccg tcc cca ccg cgt tcc caa cac gcg cac tac ctt ccc agc   385
58   Pro His Pro Ser Pro Pro Arg Ser Gln His Ala His Tyr Leu Pro Ser
59   35   40   45
60   tcg cgg cgg cca ccc agg acc ccg cgc ttc ccg ctc ccg ctg cgg atc   433
61   Ser Arg Arg Pro Pro Arg Thr Pro Arg Phe Pro Leu Pro Leu Arg Ile
62   50   55   60
63   ccc gct gcc cag cgc ccg cag gtc ctc agc acc ggg cac acg ccc ccg   481
64   Pro Ala Ala Gln Arg Pro Gln Val Leu Ser Thr Gly His Thr Pro Pro
65   65   70   75
66   acg att cca cgc cgc tgc ggg gca gga gag tcg tgg ggc aat gcc acc   529
67   Thr Ile Pro Arg Arg Cys Gly Ala Gly Glu Ser Trp Gly Asn Ala Thr
68   80   85   90   95
69   aac ctc ggc gtc ccg tgt cta cac tgg gac gag gtg ccg ccc ttc ctg   577
70   Asn Leu Gly Val Pro Cys Leu His Trp Asp Glu Val Pro Pro Phe Leu
71   100   105   110
72   gag cgg tcg ccc ccg gcc agt tgg gct gag ctg cga ggg cag ccg cac   625
73   Glu Arg Ser Pro Pro Ala Ser Trp Ala Glu Leu Arg Gly Gln Pro His
74   115   120   125
75   aac ttc tgc cgg agc ccg gat ggc tcg ggc aga cct tgg tgc ttc tat   673
76   Asn Phe Cys Arg Ser Pro Asp Gly Ser Gly Arg Pro Trp Cys Phe Tyr
77   130   135   140
78   cgg aat gcc cag ggc aaa gta gac tgg ggc tac tgc gat tgt ggt caa   721
79   Arg Asn Ala Gln Gly Lys Val Asp Trp Gly Tyr Cys Asp Cys Gly Gln
80   145   150   155
81   ggc ccg gcg ttg ccc gtc att cgc ctt gtt ggt ggg aac agt ggg cat   769
82   Gly Pro Ala Leu Pro Val Ile Arg Leu Val Gly Gly Asn Ser Gly His
83   160   165   170   175
84   gaa ggt cga gtg gag ctg tac cac gct ggc cag tgg ggg acc atc tgt   817
85   Glu Gly Arg Val Glu Leu Tyr His Ala Gly Gln Trp Gly Thr Ile Cys
86   180   185   190
87   gac gac caa tgg gac aat gca gac gca gac gtc atc tgt agg cag ctg   865
88   Asp Asp Gln Trp Asp Asn Ala Asp Ala Asp Val Ile Cys Arg Gln Leu
89   195   200   205
90   ggg ctc agt ggc att gcc aaa gca tgg cat cag gca cat ttt ggg gaa   913
91   Gly Leu Ser Gly Ile Ala Lys Ala Trp His Gln Ala His Phe Gly Glu
92   210   215   220
93   gga tct ggc cca ata ttg ttg gat gaa gta cgc tgc acc gga aac gag   961
94   Gly Ser Gly Pro Ile Leu Leu Asp Glu Val Arg Cys Thr Gly Asn Glu

```
 95            225                   230                   235
 96    ctg tca att gag caa tgt cca aag agt tcc tgg ggc gaa cat aac tgt    1009
 97    Leu Ser Ile Glu Gln Cys Pro Lys Ser Ser Trp Gly Glu His Asn Cys
 98            240                   245                   250                   255
 99    ggc cat aaa gaa gat gct gga gtg tct tgt gtt cct cta aca gat ggt    1057
100    Gly His Lys Glu Asp Ala Gly Val Ser Cys Val Pro Leu Thr Asp Gly
101                    260                   265                   270
102    gtc atc aga ctg gca gga gga aaa agt acc cat gaa ggt cgc ctg gag    1105
103    Val Ile Arg Leu Ala Gly Gly Lys Ser Thr His Glu Gly Arg Leu Glu
104                    275                   280                   285
105    gtc tac tac aag ggg cag tgg ggg aca gtc tgt gat gat ggc tgg act    1153
106    Val Tyr Tyr Lys Gly Gln Trp Gly Thr Val Cys Asp Asp Gly Trp Thr
107            290                   295                   300
108    gag atg aac aca tac gtg gct tgt cga ctg ctg gga ttt aaa tac ggc    1201
109    Glu Met Asn Thr Tyr Val Ala Cys Arg Leu Leu Gly Phe Lys Tyr Gly
110            305                   310                   315
111    aaa cag tcc tct gtg aac cat ttt gat ggc agc aac agg ccc ata tgg    1249
112    Lys Gln Ser Ser Val Asn His Phe Asp Gly Ser Asn Arg Pro Ile Trp
113            320                   325                   330                   335
114    ctg gat gac gtc agc tgc tca gga aaa gaa gtc agc ttc att cag tgt    1297
115    Leu Asp Asp Val Ser Cys Ser Gly Lys Glu Val Ser Phe Ile Gln Cys
116                    340                   345                   350
117    tcc agg aga cag tgg gga agg cat gac tgc agc cat aga gaa gat gtg    1345
118    Ser Arg Arg Gln Trp Gly Arg His Asp Cys Ser His Arg Glu Asp Val
119                    355               ·  360                   365
120    ggc ctc acc tgc tat cct gac agc gat gga cat agg ctt tct cca ggt    1393
121    Gly Leu Thr Cys Tyr Pro Asp Ser Asp Gly His Arg Leu Ser Pro Gly
122            370                   375                   380
123    ttt ccc atc aga cta gtg gat gga gag aat aag aag gaa gga cga gtg    1441
124    Phe Pro Ile Arg Leu Val Asp Gly Glu Asn Lys Lys Glu Gly Arg Val
125            385                   390                   395
126    gag gtt ttt gtc aat ggc caa tgg gga aca atc tgc gat gac gga tgg    1489
127    Glu Val Phe Val Asn Gly Gln Trp Gly Thr Ile Cys Asp Asp Gly Trp
128            400                   405                   410                   415
129    acc gat aag cat gca gct gtg atc tgc cgg cag ctt ggc tat aag ggt    1537
130    Thr Asp Lys His Ala Ala Val Ile Cys Arg Gln Leu Gly Tyr Lys Gly
131                    420                   425                   430
132    cct gcc aga gca agg act atg gct tat ttt ggg gaa gga aaa ggc ccc    1585
133    Pro Ala Arg Ala Arg Thr Met Ala Tyr Phe Gly Glu Gly Lys Gly Pro
134                    435                   440                   445
135    atc cac atg gat aat gtg aag tgc aca gga aat gag aag gcc ctg gct    1633
136    Ile His Met Asp Asn Val Lys Cys Thr Gly Asn Glu Lys Ala Leu Ala
137            450                   455                   460
138    gac tgt gtc aaa caa gac att gga agg cac aac tgc cgc cac agt gag    1681
139    Asp Cys Val Lys Gln Asp Ile Gly Arg His Asn Cys Arg His Ser Glu
140            465                   470                   475
141    gat gca gga gtc atc tgt gac tat tta gag aag aaa gca tca agt agt    1729
142    Asp Ala Gly Val Ile Cys Asp Tyr Leu Glu Lys Lys Ala Ser Ser Ser
143            480                   485                   490                   495
144    ggt aat aaa gag atg ctc tca tct gga tgt gga ctg agg tta ctg cac    1777
```

```
145        Gly Asn Lys Glu Met Leu Ser Ser Gly Cys Gly Leu Arg Leu Leu His
146                    500                 505                 510
147        cgt cgg cag aaa cgg atc att ggt ggg aac aat tct tta agg ggt gcc    1825
148        Arg Arg Gln Lys Arg Ile Ile Gly Gly Asn Asn Ser Leu Arg Gly Ala
149                    515                 520                 525
150        tgg cct tgg cag gct tcc ctc agg ctg agg tcg gcc cat gga gac ggc    1873
151        Trp Pro Trp Gln Ala Ser Leu Arg Leu Arg Ser Ala His Gly Asp Gly
152                    530                 535                 540
153        agg ctg ctt tgt gga gct acc ctt ctg agt agc tgc tgg gtc ctg aca    1921
154        Arg Leu Leu Cys Gly Ala Thr Leu Leu Ser Ser Cys Trp Val Leu Thr
155            545                 550                 555
156        gct gca cac tgc ttc aaa agg tac gga aac aac tcg agg agc tat gca    1969
157        Ala Ala His Cys Phe Lys Arg Tyr Gly Asn Asn Ser Arg Ser Tyr Ala
158        560                 565                 570                 575
159        gtt cga gtt ggg gat tat cat act ctg gta cca gag gag ttt gaa caa    2017
160        Val Arg Val Gly Asp Tyr His Thr Leu Val Pro Glu Glu Phe Glu Gln
161                    580                 585                 590
162        gaa ata ggg gtt caa cag att gtg att cac agg aac tac agg cca gac    2065
163        Glu Ile Gly Val Gln Gln Ile Val Ile His Arg Asn Tyr Arg Pro Asp
164                    595                 600                 605
165        aga agc gac tat gac att gcc ctg gtt aga ttg caa gga cca ggg gag    2113
166        Arg Ser Asp Tyr Asp Ile Ala Leu Val Arg Leu Gln Gly Pro Gly Glu
167                    610                 615                 620
168        caa tgt gcc aga cta agc acc cac gtt ttg cca gcc tgt tta cct cta    2161
169        Gln Cys Ala Arg Leu Ser Thr His Val Leu Pro Ala Cys Leu Pro Leu
170            625                 630                 635
171        tgg aga gag agg cca cag aaa aca gcc tcc aac tgt cac ata aca gga    2209
172        Trp Arg Glu Arg Pro Gln Lys Thr Ala Ser Asn Cys His Ile Thr Gly
173        640                 645                 650                 655
174        tgg gga gac aca ggt cgt gcc tac tca aga act cta caa caa gct gct    2257
175        Trp Gly Asp Thr Gly Arg Ala Tyr Ser Arg Thr Leu Gln Gln Ala Ala
176                    660                 665                 670
177        gtg cct ctg tta ccc aag agg ttt tgt aaa gag agg tac aag gga cta    2305
178        Val Pro Leu Leu Pro Lys Arg Phe Cys Lys Glu Arg Tyr Lys Gly Leu
179                    675                 680                 685
180        ttt act ggg aga atg ctc tgt gct ggg aac ctc caa gaa gac aac cgt    2353
181        Phe Thr Gly Arg Met Leu Cys Ala Gly Asn Leu Gln Glu Asp Asn Arg
182                    690                 695                 700
183        gtg gac agc tgc cag gga gac agt gga gga cca ctc atg tgt gaa aag    2401
184        Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Glu Lys
185            705                 710                 715
186        cct gat gag tcc tgg gtt gtg tat ggg gtg act tcc tgg ggg tat gga    2449
187        Pro Asp Glu Ser Trp Val Val Tyr Gly Val Thr Ser Trp Gly Tyr Gly
188        720                 725                 730                 735
189        tgt gga gtc aaa gac act cct gga gtt tat acc aga gtc ccc gcc ttt    2497
190        Cys Gly Val Lys Asp Thr Pro Gly Val Tyr Thr Arg Val Pro Ala Phe
191                    740                 745                 750
192        gta cct tgg ata aaa agt gtc acc agt ctg taacttatgg aaagctcaag      2547
193        Val Pro Trp Ile Lys Ser Val Thr Ser Leu
194                    755                 760
```

```
195          aaaatagtaa aacagtaacc attcagtctt catacttggc accatgccag aaaaaaaaaa      2607
196          aaaaaaa                                                                 2614
197   <210> SEQ ID NO 4
198   <211> LENGTH: 761
199   <212> TYPE: PRT
200   <213> ORGANISM: Mouse
201   <220> FEATURE:
202   <223> OTHER INFORMATION:
203   <400> SEQUENCE: 4
204          Met Ala Leu Ala Arg Cys Val Leu Ala Val Ile Leu Gly Ala Leu Ser
205           1               5                  10                  15
206          Val Val Ala Arg Ala Asp Pro Val Ser Arg Ser Pro Leu His Arg Pro
207                      20                  25                  30
208          His Pro Ser Pro Pro Arg Ser Gln His Ala His Tyr Leu Pro Ser Ser
209                      35                  40                  45
210          Arg Arg Pro Pro Arg Thr Pro Arg Phe Pro Leu Pro Leu Arg Ile Pro
211                  50                  55                  60
212          Ala Ala Gln Arg Pro Gln Val Leu Ser Thr Gly His Thr Pro Pro Thr
213          65                  70                  75                  80
214          Ile Pro Arg Arg Cys Gly Ala Gly Glu Ser Trp Gly Asn Ala Thr Asn
215                          85                  90                  95
216          Leu Gly Val Pro Cys Leu His Trp Asp Glu Val Pro Pro Phe Leu Glu
217                      100                 105                 110
218          Arg Ser Pro Pro Ala Ser Trp Ala Glu Leu Arg Gly Gln Pro His Asn
219                      115                 120                 125
220          Phe Cys Arg Ser Pro Asp Gly Ser Gly Arg Pro Trp Cys Phe Tyr Arg
221              130                 135                 140
222          Asn Ala Gln Gly Lys Val Asp Trp Gly Tyr Cys Asp Cys Gly Gln Gly
223          145                 150                 155                 160
224          Pro Ala Leu Pro Val Ile Arg Leu Val Gly Gly Asn Ser Gly His Glu
225                      165                 170                 175
226          Gly Arg Val Glu Leu Tyr His Ala Gly Gln Trp Gly Thr Ile Cys Asp
227                      180                 185                 190
228          Asp Gln Trp Asp Asn Ala Asp Ala Asp Val Ile Cys Arg Gln Leu Gly
229                  195                 200                 205
230          Leu Ser Gly Ile Ala Lys Ala Trp His Gln Ala His Phe Gly Glu Gly
231              210                 215                 220
232          Ser Gly Pro Ile Leu Leu Asp Glu Val Arg Cys Thr Gly Asn Glu Leu
233          225                 230                 235                 240
234          Ser Ile Glu Gln Cys Pro Lys Ser Ser Trp Gly Glu His Asn Cys Gly
235                          245                 250                 255
236          His Lys Glu Asp Ala Gly Val Ser Cys Val Pro Leu Thr Asp Gly Val
237                      260                 265                 270
238          Ile Arg Leu Ala Gly Gly Lys Ser Thr His Glu Gly Arg Leu Glu Val
239                  275                 280                 285
240          Tyr Tyr Lys Gly Gln Trp Gly Thr Val Cys Asp Asp Gly Trp Thr Glu
241              290                 295                 300
242          Met Asn Thr Tyr Val Ala Cys Arg Leu Leu Gly Phe Lys Tyr Gly Lys
243          305                 310                 315                 320
244          Gln Ser Ser Val Asn His Phe Asp Gly Ser Asn Arg Pro Ile Trp Leu
```

Line	?	Error/Warning	Original Text
26	W	"N" or "Xaa" used: Feature required	gtgctcacng cngcbcaytg
39	W	"N" or "Xaa" used: Feature required	agcggnccnc cdgartcvcc